

깊이 터치를 통한 영상 이벤트 제어 시스템

(Video event control system by recognition of depth touch)

이 동 석¹⁾, 권 순 각^{2)*}

(Dong-Seok Lee and Soon-Kak Kwon)

요 약 재생되는 동영상에서 정지, 재생, 캡처, 확대, 축소 등의 다양한 이벤트 제공은 스마트폰과 같은 작은 크기의 모니터에서 가능하다. 그러나, 모니터의 크기가 커지게 되면, 터치 인식에 대한 비용이 증가하게 되어, 터치를 통한 이벤트 제공은 사실상 불가능하다. 본 논문에서는 저렴하게 깊이 정보로 터치를 인식하고, 단일 및 다중 터치로 토글, 핀치-인/아웃 등의 다양한 이벤트를 부여하는 영상 이벤트 제어 시스템을 제안한다. 깊이 카메라로부터 얻어진 깊이 정보로 터치된 위치와 터치 경로를 찾고, 터치 제스처 종류를 파악한다. 이러한 터치 인터페이스 알고리즘은 소형 싱글보드 시스템에서 구현하고, UART 통신을 통해 제스처 정보를 전송함으로써 영상 이벤트를 제어할 수 있다. 모의실험을 바탕으로 대형 스크린에서 제안한 깊이 터치 방법으로 다양한 영상 이벤트를 제어할 수 있음을 보인다.

핵심주제어 : 깊이 정보, 터치 제스처, 터치 인터페이스, 이벤트 제어

Abstract Various events of stop, playback, capture, and zoom-in/out in playing video is available in the monitor of a small size such as smart phones. However, if the size of the display increases, then the cost of the touch recognition is increased, thus provision of a touch event is not possible in practice. In this paper, we propose a video event control system that recognizes a touch inexpensively from the depth information, then provides the variety of events of the toggle, the pinch-in / out by the single or multi-touch. The proposed method finds a touch location and the touch path by the depth information from a depth camera, and determines the touch gesture type. This touch interface algorithm is implemented in a small single-board system, and can control the video event by sending the gesture information through the UART communication. Simulation results show that the proposed depth touch method can control the various events in a large screen.

Key Words : Depth information, Touch gesture, Touch interface, Event control

1. 서 론

영상 이벤트 제어는 동영상에 대해 재생/일시 멈춤, 특정 화면의 캡처, 캡처된 화면의 크기 확대/축소 등 영상의 재생에 대해 여러 가지 이벤트를 제어하기 위하여 사용된다. 모바일 기기에서는 영상 이벤트를 제어하기 위해서 터치 인터페이스가 널리 적용되고 있다. 최근 터치 기술은 기존 키보드나 마우스, 키패드 등의 물리적 인터

* Corresponding Author : skkwon@deu.ac.kr

† 본 연구(No. C0298779)는 중소기업청에서 지원하는 2015년도 산학연협력 기술개발사업 연구수행으로 인한 결과물임을 밝힙니다.
Manuscript received February 12, 2016 / accepted February 19, 2016

1) 동의대학교 컴퓨터소프트웨어공학과, 제1저자

2) 동의대학교 컴퓨터소프트웨어공학과, 교신저자

페이스를 모니터에 직접 터치하는 인터페이스로 대체하는 기술로 지속적으로 발전해 오고 있다. 실제로 터치를 이용한 인터페이스의 경우 기존 물리적 인터페이스를 사용할 경우보다 입력 속도가 2배 이상 빠르다[1-2]. 또한 터치 인터페이스는 성별에 관계없이 편리하게 이용할 수 있고, 인터페이스를 숙달하는데 기존 방법보다 빠르게 숙달이 가능하다. 이러한 이점으로 인해 일반 사용자를 대상으로 한 제품뿐만 아니라 여러 산업 분야에서도 터치 인터페이스가 기존의 물리적 인터페이스를 대체해가고 있는 실정이다. 또한 터치를 이용한 응용 분야에 대한 연구도 계속 이루어지고 있다[3-5].

하지만 대형 스크린 환경에서는 현재까지는 터치 인터페이스를 적용하기가 어렵다. 대형 스크린 환경에서 기존 물리적 터치스크린을 설치하여 터치 인터페이스를 적용하는 방법은 막대한 비용이 든다. 이러한 문제점을 해결하고자 스크린을 타고 흐르는 적외선 신호를 송수신하여 이를 감지하여 터치를 인식하는 방법[6-7], 초음파를 발사하는 펜을 이용하여 그 초음파를 인식하여 터치를 인식하는 방법[8-9] 등이 제안되기도 하였다. 하지만 이러한 방법들은 초기 설정이 어렵거나 별도의 도구가 필요한 등의 문제점이 있다.

본 논문에서는 깊이 터치를 이용해 여러 개의 터치를 인식하고, 각각의 터치를 추적하여 제스처를 판단하여 그 제스처에 따라서 제어를 할 수 있는 터치 인터페이스를 구현한다. 깊이 터치는 깊이 카메라를 이용하여 깊이 영상을 촬영하고, 깊이 영상내의 깊이 값을 이용하여 터치를 인식하는 기술이다[10-11]. 깊이 터치 기술을 이용하여 터치를 인식할 수 있을 뿐만 아니라 터치 점을 추적하여 여러 제스처를 인식할 수 있다. 이러한 알고리즘을 소형화된 싱글보드 기기에 탑재하여 제스처를 신속하게 인식할 수 있으며, UART 방식을 통하여 제어시스템과 연결하여 여러 이벤트를 제공할 수 있다. 이러한 방식을 통해 저렴한 가격으로 대형 스크린 환경에서 영상 이벤트 제어를 수행할 수 있다. 또한 새로운 기능을 추가함에 있어서도 기존 시스템의 변경 없이 제스처 인식 부분만 소프트웨어 수정 및 교체를 통해서 가능하기 때문에 확장성이 높다. 본

논문에서 제안된 방법을 통해 터치 인터페이스 시장에 대해 대형 스크린 환경에서의 영상 이벤트 분야로 확장해 나갈 수 있다.

2. 시스템 구성

2.1 깊이 터치를 통한 영상 이벤트 제어 시스템

Fig. 1은 깊이 카메라를 이용하여 깊이 터치를 통한 영상 이벤트 제어 시스템을 구현하기 위한 시스템 구성도이다. 깊이 카메라는 깊이 영상을 촬영하여 촬영된 깊이 영상을 전송한다. 촬영된 깊이 영상을 이용하여 깊이 터치를 인식하고, 인식된 깊이 터치를 추적하여 제스처를 판단한다. 제스처가 판단되면 UART 통신을 통해 영상 이벤트 제어 시스템에 제스처 종류를 전송하여 영상 이벤트 제어 시스템이 알맞은 이벤트를 수행하여 영상을 제어한다.

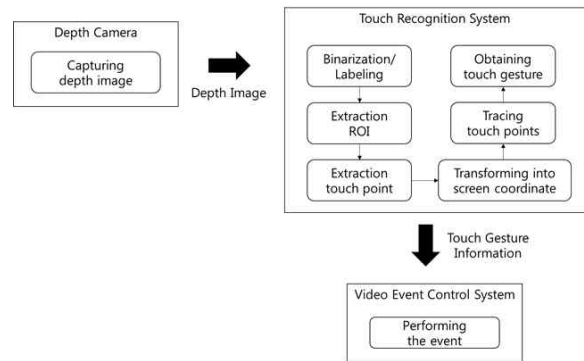


Fig. 1 Flowchart of the proposed system.

본 논문에서 깊이 터치를 통한 영상 이벤트 제어 시스템을 구현하기 위해 깊이 카메라로는 ASUS사의 Xtion Pro를, 터치를 인식하는 알고리즘을 수행하는 소형 싱글보드 시스템은 HardKernel사의 ODroid-C1를 사용하였다. ODroid-C1은 ARM 프로세서를 이용하며, USB 인터페이스와 GPIO 등 여러 인터페이스를 지원하는 보드이다. 또한 ODroid-C1은 Linux, Android 등의 여러 범용 OS를 지원하기 때문에 개발에 있어서도 C나

C++등의 보편적인 프로그래밍 언어를 이용할 수 있다는 장점이 있다. 또한 UART 킷을 제공하므로, 이를 이용하여 제스처 정보를 신속하게 전송할 수 있다.

2.2 깊이 터치 제스처 인식 알고리즘

깊이 카메라를 이용하여 배경 영상을 찾아내고, 촬영된 깊이 영상을 배경 영상과 비교하는 방법을 통해 깊이 터치를 인식할 수 있다[10-11]. 또한 깊이 터치를 추적하여 터치 제스처를 인식할 수 있다.

깊이 카메라를 통해 깊이 영상을 촬영하여 깊이 영상을 얻는다. 깊이 영상에서 각각의 화소에 대하여 촬영된 깊이 영상의 깊이 값과 배경 깊이 영상의 깊이 값을 비교하여 차이가 많이 나는 지점에 대해 이진화를 수행한 후[11] 이진화 과정에서 분류된 객체와 배경 깊이 값을 비교하여 배경 깊이 값과 차이가 크지 않는 지점을 깊이 터치로 인식한다.

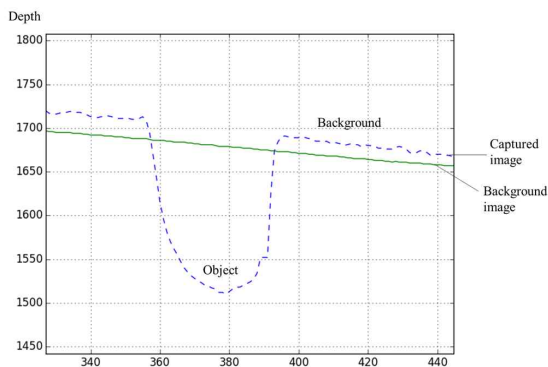


Fig. 2 Method of the binarization by the depth image.

촬영 영상에서의 좌표계는 실제 이용하고자 하는 스크린의 좌표계는 서로 다르다. 따라서 실제로 터치를 이용하여 위치를 이용한 이벤트를 제공하기 위해서는 영상 좌표계를 스크린의 좌표계로 변환하는 캘리브레이션 과정이 초기화 단계에서 수행되어야 한다. 이때 깊이 영상에서의 스크린 평면은 원근 왜곡이 일어나 Fig. 3처럼 기하학적인 변형이 일어나게 되어 사다리 형태로 촬

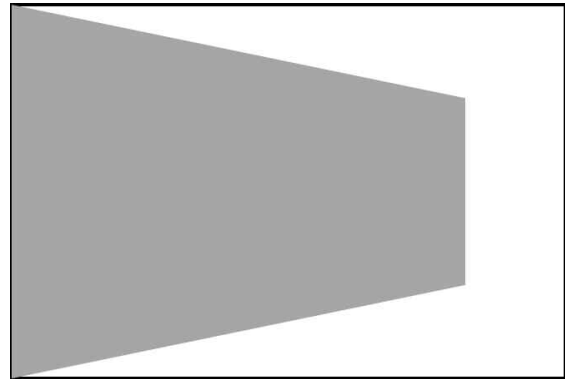


Fig. 3 Screen shape in the captured image.

영된다. 이는 촬영된 스크린의 영상이 하나의 평면형 물체로 간주할 수 있다는 것을 의미한다. 따라서 식 (1)과 같은 호모그래피 변환을 통해 스크린 상의 좌표계로 전환할 수 있다. 여기서 H는 호모그래픽으로 표현된 3X3 행렬이며, w는 스케일을 의미한다. H, w는 두 좌표계의 매칭되는 좌표들의 쌍들을 DLT(Direct Linear Transform) 식에 적용하여 호모그래피를 추정하여 구할 수 있다[12]. 또한 x, y는 변환 전의 좌표계에서의 좌표이고, x', y'는 변환 후 좌표이다.

$$w \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (1)$$

Table 1 Supported touch gesture in the proposed touch interface.

Touch gesture	Action
Single touch	
Two touch	
Drag	
Pinch-in	
Pinch-out	

터치된 점들에 대해 터치 경로와 스크린 상의 위치정보를 알 수 있다면 제스처의 종류를 알 수

있다. 제안하는 터치 인터페이스 시스템은 기존에 널리 사용되어 사용자에게 익숙한 터치, 멀티 터치, 드래그, 핀치 인/아웃을 제공한다.

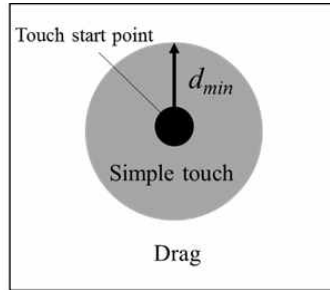


Fig. 4 Distinguishing gesture between the simple touch and the drag.

이러한 제스처를 지원하기 위해서는 먼저 터치가 단순한 터치인지 드래그 상황인지를 구분해야 한다. 단순한 터치와 드래그 제스처를 구분하기 위해 터치 경로를 추적하여 그 이동거리가 d_{min} 이하로 움직인다면 단순한 터치로 인지한다. 반면 터치 이동거리가 최소 제스처 거리 이상이라면 그 터치에 대해 드래그 제스처로 판단한다.

또한 두개의 터치가 인식되었다면 멀티 터치 혹은 핀치 인/아웃 제스처를 적용한다. 먼저 두 점에 대해 각각을 드래그 상태인지 혹은 단순 터치인지 판단할 수 있다. 만약 두 점 다 터치 상태이면 멀티 터치이다. 반면 한 점이라도 드래그 상태이라면 핀치 인/아웃 제스처라고 판단한다. 그 후 두 터치점의 시작 점과 마지막 점을 찾아낸다. 두 터치점의 각 시작점간 거리 d_s 와 각 마지막 점간 거리 d_e 를 측정하여 d_s 와 d_e 의 크기

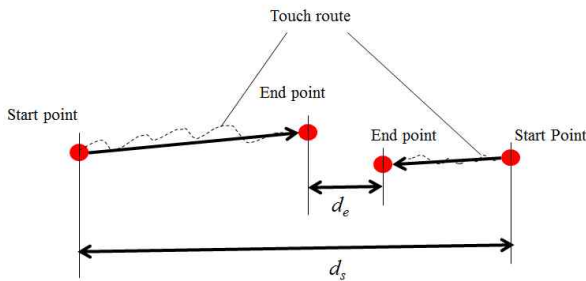


Fig. 5 Distinguishing the pinch-in and pinch-out gestures.

를 비교한다. 만약 d_s 보다 d_e 가 크다면 이 터치는 핀치 아웃 제스처로 판단한다. 반대의 상황이라면 핀치 인 제스처로 판단한다.

제스처 종류가 구분되었다면 제스처 정보를 UART 통신 방법을 통하여 영상 이벤트 제어 시스템에 전송한다. UART통신은 RS-232, RS-422, RS-485등의 통신 인터페이스 표준을 적용한 통신 방법이다. UART는 메모리 또는 레지스터에 존재하는 데이터를 한 패킷에 5-8비트 단위로 직렬화하여 시작 비트와 패리티 비트, 종료비트를 추가하여 비동기적으로 통신한다. Fig. 6은 UART 통신을 할 때의 프로토콜이다. 이 때 통신하고자 하는 양쪽에서 통신 속도를 미리 약속하고 같은 속도로 설정해야 통신이 원활하게 이루어진다.

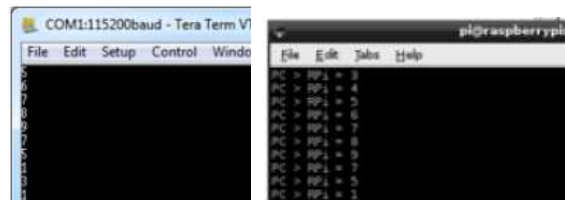
Bit number	1	2	3	4	5	6	7	8	9	10	11
	Start bit	5-8 data bits								Parity bit	Stop bit
	Start	Data 0	Data 1	Data 2	Data 3	Data 4	Data 5	Data 6	Data 7	Parity	Stop

Fig. 6 The data framing of UART communication

제안하는 시스템에서 사용하는 소형 싱글보드 기기인 ODroid는 UART 통신을 USB 포트를 통해 연결할 수 있게 하는 키를 제공한다. 이를 이



(a)



(b)

Fig. 7 UART communicate method: (a) UART kit for the ODroid single-board system series, (b) UART communicate example.

용하여 간편하게 UART 통신을 수행할 수 있다. UART 통신을 통하여 제스처 종류를 전송할 때 먼저 제스처 종류를 전송한 후 그 후 각 터치점의 스크린 좌표 정보를 전송한다.

2.3 깊이 터치를 통한 영상 이벤트 제어 시스템 구현

본 논문에서는 제안한 방법을 이용하여 깊이 터치를 통한 영상 이벤트 제어 시스템을 구현하였다. 이 시스템은 미리 녹화된 영상에 대해 재생/일시멈춤/화면 캡처를 지원하고, 캡처된 화면에 대해 확대, 축소를 지원하는 시스템이다.

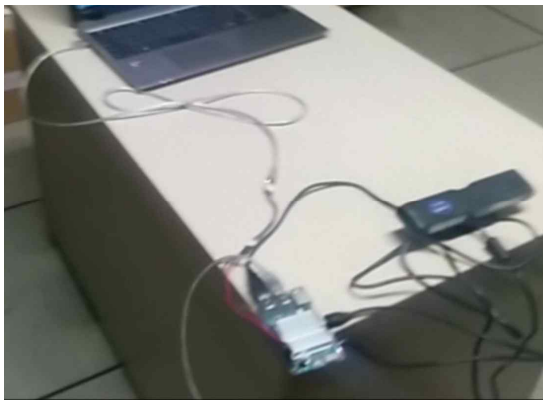


Fig. 8. Example of the proposed system.

먼저 스크린을 촬영할 수 있는 곳에 깊이 카메라를 설치하고, 터치 인터페이스 시스템이 탑재된 소형 싱글보드 시스템에 깊이카메라를 연결한다. 또한 UART 키를 싱글보드 시스템의 GPIO 핀과 연결하고 UART 키의 USB 포트를 영상 이벤트 제어시스템에 연결한다.

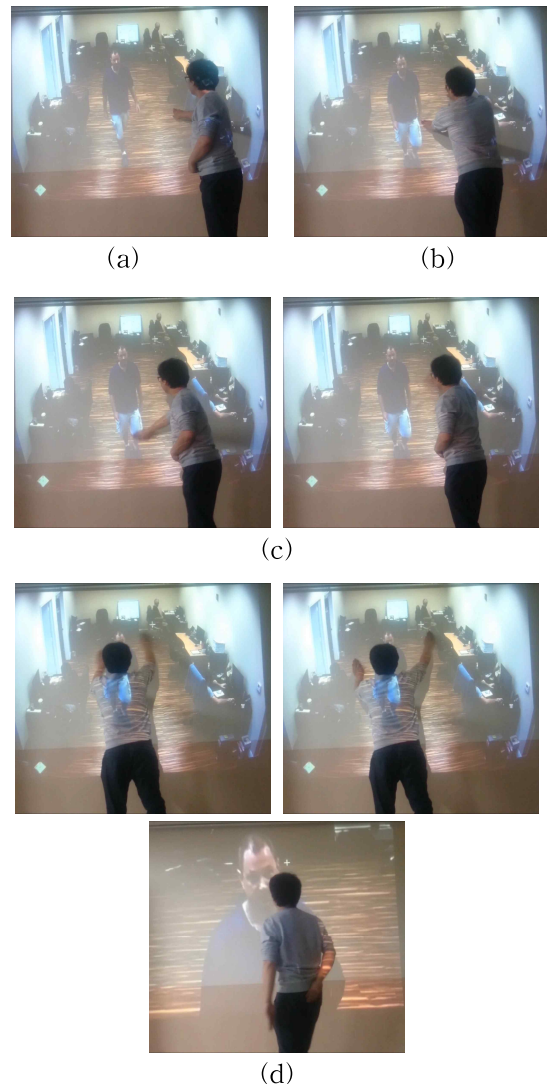
구현된 영상 이벤트 제어 시스템에서 에서 깊이 터치 제스처의 종류에 대응하여 수행하는 이벤트는 Table 2와 같다. 이 때 핀치 인/아웃 제스처를 통해 확대/축소 이벤트를 수행할 때는 두 터치 점의 중심을 중앙으로 영상을 확대/축소한다.

Fig. 9는 빔 프로젝터를 이용한 대형 스크린 환경에서 깊이 터치 제스처를 인식하여 영상 제어를 하는 장면이다. 이 때 대형 스크린에서

Table 2 Events corresponding with the type of gestures.

Touch gesture	Event
Single touch	Play/Stop of the video (toggle)
Two touch	Getting a snapshot of current playing video
Drag	Showing the snapshot image
Pinch-out	Zoom-in within the current image
Pinch-in	Zoom-out within the current image

터치 제스처 인식이 원활하게 이루어져서 영상 이벤트를 제어할 수 있음을 확인하였다.



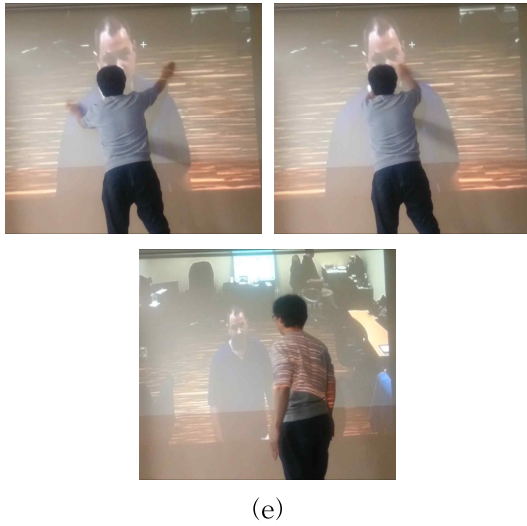


Fig. 9 Examples of controlling video event by depth touch gestures: (a) single touch, (b) two touch, (c) drag, (d) pinch-out, (e) pinch-in.

이러한 터치 제스처에 대해 이벤트가 처리되는 반응속도를 측정하였다. 반응속도가 눈에 띄게 느리다면 사용자는 눈에 띄는 불편함을 느끼게 되기 때문에 사용자의 편의성에 있어서 중요한 성능 척도이다. 이 때 실험에 사용한 소형 싱글 보드 시스템인 ODroid-C1의 사양은 Table 3과 같고, 깊이터치 제스처 정보를 받아 영상을 제어하는 시스템의 CPU는 AMD Dual Core Processor E-450 제품으로써 최대 CPU 클럭이 1.65Ghz이다.

Table 3 Specification of the ODroid-C1

CPU	Amlogic ARM@ Cortex@-A5(ARMv7) 1.5Ghz quad core CPUs
GPU	MaliTM-450 MP2 GPU (OpenGL ES 2.0/1.1 enabled for Linux and Android)
RAM	1Gbyte DDR3 SDRAM
Network	Gigabit Ethernet

실험은 각 제스처에 대해 각각 20회 반복하여

제스처를 수행한 시간부터 이벤트가 수행된 시간까지를 측정하였다. 실험 결과 터치 제스처에 대한 반응 속도가 평균 0.297초가 소요되는 것으로 나타났다. 이는 사용자가 사용하기에 체감상 거의 시간차이를 느끼지 못하는 수치로서, 제안된 시스템을 통해 인터페이스를 편리하게 이용할 수 있다는 것을 의미한다.

Table 4 The result of the reaction time in the proposed system.

Touch gesture	The reaction speed(s)
Single touch	0.293
Two touch	0.295
Drag	0.298
Pinch-out	0.302
Pinch-in	0.301

3. 결론

본 논문에서는 깊이 정보를 이용하여 깊이 터치를 인식하고, 터치 제스처를 파악하여 이를 통해 영상 이벤트를 제어하는 시스템을 구현하였다. 영상 이벤트는 재생되는 동영상에서 특정 위치의 화면을 선택하고, 확대하는 핀치 기능이 중요하며, 이를 위해서는 재생되는 동영상의 일시 정지 기능, 다시 재생하는 기능, 특정 화면을 캡처하는 기능, 캡처된 화면을 확대하거나 축소하는 기능 등이 필요하다. 본 논문에서는 소형뿐만 아니라 중대형 크기의 디스플레이 모니터에서도 이러한 영상 이벤트 제어를 제공해주기 위하여 깊이 터치를 통해 터치 제스처 인터페이스를 구현하는 방법을 제안하였다. 기존 물리적 터치스크린을 도입하는 방법에 비해 훨씬 저렴하게 터치 제스처 인터페이스를 대형의 스크린에 적용하였고, 기존 물리적 인터페이스를 이용한 방법에 비해 조작이 쉽고 간편하여 인터페이스를 숙달하는데 짧은 시간밖에 들지 않는다. 본 논문에서

제안한 깊이 터치를 통한 영상 이벤트 제어 방법은 중대형 스크린을 이용한 관제 제어 시스템, 교육[13] 등의 스크린 환경으로 응용될 수 있을 것으로 기대된다.

References

- [1] K. Kin, M. Agrawala and T. DeRose, "Determining the benefits of direct-touch, bimanual, and multifinger input on a multitouch workstation," *Proceeding of Graphics Interface 2009*, pp. 119-124, 2009.
- [2] S. B. Han, S. Y. Yoon and Tae-Kyun Kim, "A study on the application of design through comparative assessment of users' experience in regard to the input interface mode on mobile phones," *Journal of Digital Design*, Vol. 11, No. 1, pp. 325-334, 2011.
- [3] D. H. Lee, M. K. Kang and T. S. Yun, "Development of a tangible interface using multi-touch display on an irregular surface," *Journal of the Korea Industrial Information Systems Research*, Vol. 16, No. 3, pp. 65-72, 2011.
- [4] J. I. Lee, S. C. Kim and H.C. Won, "Voice-based messenger using NXT touch-sensor input unit and the bluetooth wireless communication for the blind," *Journal of the Korea Industrial Information Systems Research*, Vol. 13, No. 5, pp. 78-86, 2008.
- [5] E. S. Ko, Y. W. Rhee and C. W. Lee, "Implementation of commercial IWB interface using image processing," *Journal of the Korea Industrial Information Systems Research*, Vol. 17, No. 6, pp. 19-24, 2012.
- [6] V. Soni, M. Patel and R. S. Narde, "An interactive infrared sensor based multi-touch panel," *International Journal of Scientific and Research Publications*, Vol. 3, No. 3, 2013.
- [7] S. W. Jung, O. J. Kwon and Y. J. Jeong, "Design of an infrared multi-touch screen controller using stereo vision," *The Institute of Electronics Engineers of Korea - Semiconductor and Devices*, Vol. 47, No. 2, pp. 68-76, 2010.
- [8] H. Nonaka and T. Da-te, "Ultrasonic position measurement and its applications to human interface," *IEEE Transactions on Instrumentation and Measurement*, Vol. 44, No. 3, pp. 771-774, 1995.
- [9] G.F. Russell, B.A. Smith and T.G. Zimmerman, *Digital Pen using Ultrasonic Tracking*. 6703570, U.S., 2004.
- [10] D. S. Lee and S. K. Kwon, "Touch pen using depth information," *Journal of Korea Multimedia Society*, Vol. 18, No. 11, pp. 1313-1318, 2015.
- [11] D. S. Lee and S. K. Kwon, "Touch event method using depth information," *Proceeding of The 11th International Conference on Multimedia Information Technology and Applications*, 2015.
- [12] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge university Press, Cambridge, 2004.
- [13] S. K. Kwon, Y. H. Park, O. J. Kwon, and S. W. Han, "Cascade mentoring system for computer major education," *Journal of the Korea Industrial Information Systems Research*, Vol. 20, No. 5, pp. 71-80, 2015.



이 동 석 (Dong-Seok Lee)

- 정회원
- 동의대학교 컴퓨터소프트웨어 공학과 공학사
- 동의대학교 컴퓨터소프트웨어 공학과 석사과정
- 관심분야 : 멀티미디어 신호처리, 영상 인식



권 순 각 (Soon-Kak Kwon)

- 정회원
- 경북대학교 전자공학과 공학사
- KAIST 전기및전자공학과 공학 석사
- KAIST 전기및전자공학과 공학박사
- 동의대학교 컴퓨터소프트웨어공학과 교수
- 관심분야 : 멀티미디어처리, IOT